DOCUMENT RESUME

ED 130 096 CE 008 559

AUTHOR Brannon, Donald R.
TITLE Issues in Vocational Education Evaluation.
PUB DATE Apr 76
NOTE 18p.; Paper presented at the Southern Region Vocational Education Leadership Development Conference (Atlanta, Georgia, April 4-7, 1976). For related papers see CE 008 557-570

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Policy; Educational Problems; *Educational Trends; Efficiency; Evaluation; Program Effectiveness; *Vocational Education

ABSTRACT

One of a series of papers on critical issues in vocational-technical education, this paper identifies and explores four issues in evaluation of vocational education: (1) Time and Performance, (2) Evaluation of Outputs and/or Outcomes, (3) Evaluation of Efficiency and Effectiveness, and (4) Policy Evaluation. (HD)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED130096

ISSUES IN VOCATIONAL EDUCATION EVALUATION

Donald R. Brannon

Department of Occupational Education
School of Education
North Carolina State University
Raleigh, North Carolina

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

April, 1976

CE 001 559

ED130096

Introduction

Evaluation in vocational and other, complementary forms of education has changed over time just as have other aspects of the American society. There are as many proponents of differing forms of evaluation as there are differing definitions of evaluation. It is not my purpose here to advocate one form of evaluation over another; it is my purpose to discuss issues which, in my judgment, transcend evaluation of performance in vocational education.

However, for the purposes of this paper, I must use a definition of evaluation. It is not the definition of evaluation--just a definition of evaluation: Evaluation is the process of furnishing decision-makers with information about an entity or state of affairs which decision-makers can use to judge the worth or value of the entity or state of affairs. This definition is, in a sense, a policy statement about evaluation, and it sets the parameters within which I shall operate. Some assumptions underlie this definition, and it is within this definition that we shall explore some issues regarding evaluation in vocational education.

Before we discuss issues, we should have some common ground on what an issue is. For the purposes of this paper, an issue is the difference between "what is" and "what ought to be" that cannot be reconciled by the use of only empirical or factual means. An issue is essentially a value question, the answer to which depends on the values of each decision-maker. Thus, the issues discussed herein concerning evaluation in vocational education obviously reflect my own values. The issues to be discussed are listed here in order of presentation:

1. Time and Performance
2. Evaluation of Outputs and/or Outcomes
3. Evaluation of Efficiency and Effectiveness
4. Policy Evaluation

Again, it should be noted that the following discussion is not aimed at providing answers. Answers to issues depend upon the readers' values. You must help to decide the answers.

Time and Performance

The first issue deals with the relative importance of time and performance in assessing the ability of students to perform in an occupational situation. Presently, there are three main quantifiable methods of evaluating students of vocational education: (1) hours completed, (2) the use of norm-referenced tests, and (3) the use of criterion-referenced tests. Time has traditionally been the major variable of interest in gauging the progress of students toward reaching an acceptable level of occupational performance. Apprenticeships, for example, use primarily time-referenced criteria. Regardless of how proficient the apprentice is at the beginning of training, the length of the training period seldom varies. As another example, many vocational-technical schools quote the number of hours of instruction as evidence of achieved performance.

CE 008 559

An alternative to the time variable is the performance variable, but even here a significant issue of absolute standards versus relative standards has developed during the past decade. Norm-referenced tests basically advocate relative standards--the relative performance of student X in relation to his or her peers. In recent years, the relative standard has been questioned and is beginning to be supplanted by the absolute standard, which has given rise to the criterion-referenced test. Which of these is most accurate? For evaluating what? As a means of contrasting these methods, I will discuss hours completed first and then compare the usage of criterion-referenced and norm-referenced tests.

The history of the use of the number of hours completed in learning as a means of evaluating competencies goes back at least to the guild system. This usage was formally institutionalized in the United States more than 50 years ago, about 1916, and has since been referred to as use of the Carnegie unit. Use of the Carnegie unit for evaluative purposes of competencies carries with it at least one assumption--specifically, that the number of hours completed while working on becoming competent represents a valid indicator of achieved competency. Whether or not this assumption is correct and the degree to which it may be correct are questions that vocational educators must explore. In some labor apprenticeship programs, completion of a certain number of hours changes a person from an apprentice into a journeyman. For example, some plumbers' unions require six years of apprenticeship before an apprentice becomes a journeyman. This length of time is supposed to assure highly developed competencies; perhaps more significantly, this also helps to assure higher plumbers' wages because fewer people have access to the occupational field. The same principle holds true in education. We may find that many diplomas at all education levels are given primarily as certificates of hours completed. Since many teaching fields are becoming overcrowded, another year of education may be added to the present four-year baccalaureate requirement. This extra year might both improve teacher competencies and alleviate some of the overcrowding problems. The more overloaded the occupational area is (or the more desire there is for higher wages), the more hours may be required in order to qualify for work in that area. If an occupational field becomes too overcrowded, wages may decrease because buyers of competencies have more people from whom to select and can bargain down the price of competency. According to one study, the real monetary value of a college education decreased by eleven per cent between 1969 and 1975 (Freeman and Holloman, 1975).

If, however, total hours completed in a field <u>may not</u> validate competency development in that field, the number of hours completed may have a purpose beyond the evaluative purpose we commonly use. If the prime goal of using total hours completed is controlling the job market supply, then how can we realistically evaluate students as if total hours primarily indicated competency? This use of hours completed as an evaluative tool may be analogous to using a sledge hammer to evaluate how easily different kinds of tacks can be driven into wood. Some tacks may go in more easily than others when hit by a sledge hammer, but what realistic carpenter would use a sledge hammer to drive tacks into wood? The carpenter would use a sledge hammer if the choice were between the sledge hammer and his hand, but not

if the choice were between a sledge hammer and a tack hammer. Correspondingly, each evaluator needs to find out if there are evaluative methods that more closely resemble the precision gained in the use of tack hammers for driving tacks.

Two methods whose use should be more accurate for judging students' performance than the use of hours completed are (1) criterion-referenced tests and (2) norm-referenced tests. Chart A is a comparison of the uses of criterion-referenced and norm-referenced tests.

CHART A

A Contrast Between Relevant Aspects of Criterion-Referenced Tests and Norm-Referenced Tests

Aspects of Criterion-Referenced Tests	Aspects of Norm-Referenced Tests
1. [Evaluate] the content of behavioral repertory (Glaser, 1963).	1. [Evaluate] if one student is more or less proficient than another (Glaser, 1963).
2. [Evaluate] the correspondence between what an individual does and the underlying continuum of achievement (Glaser, 1963).	2. Do not tell how proficient students are with respect to the subject matter tasks involved (Glaser, 1963).
3. Used to evaluate an individual's status with respect to some criterion, i.e., performance standard (Popham and Husek, 1969).	3. Used to evaluate an individual's performance in relation to the performance of other individuals on the same measuring device (Popham and Husek, 1969).
4. Used to make decisions about individuals and treatments, e.g., instructional sequences (Shoemaker, 1971).	4. Used to make decisions about comparisons between individuals, e.g., situations requiring selectivity (Popham and Husek, 1969).
5. Use item analysis which discriminates between pre- and posttest groups and seems appropriate (Cox, 1970).	5. Use item analysis which selects items that maximize differences among individuals and seems appropriate, e.g., the concept of discrimination (Cox and Vargas, 1965).
6. Related closely to the development of individualized instruction and testing (Davis, 1972).	6. Related closely to the development of group instruction and testing.

7. Not designed to produce a wide range of scores.	7. Designed to produce a wide range of scores (Klein and Kosecoff).
8. Content validity is overriding (Gagne, 1969).	8. Predictive validity is overriding (Gagne, 1969).
9. Student competes with present criteria and his own past record (Thorndike, 1913).	9. Student competes with other students.
10. Scores indicate specific behaviors that the individual masters or does not master, without regard to other individuals (Cox and Vargas, 1967).	10. Scores indicate the relative standing of an individual in a group (Popham and Husek, 1969).
11. Utility comes from the meaningfulness and usefulness the information has for formulating instructional sequences (Shoemaker, 1971).	11. Utility comes from convenience of handling test score data, equality of units, the equating of forms, and comparing individual scores on the forms (Angoff, 1962).
12. Emphasis is on an absolute level of performance (Klein and Kosecoff).	12. Emphasis is on a relative level of performance (Klein and Kosecoff).

Some limitations of these two methods follow: For evaluative purposes criterion-referenced tests aid in determining the degree of student accomplishment with respect to some criteria. Using only criterion-referenced scores, one does not compare one student with others. For example, if one were trying to determine how competent a person is in a skill area, then a criterion-referenced test would seem appropriate to use. However, if one wanted to compare student scores with those of others, then one would use norm-referenced tests. In my opinion, criterion-referenced tests are more appropriate for evaluating vocational education students than are norm-referenced tests, for the following reasons.

More and more, questions are being raised about which groups norm-referenced tests are normed for. Most questions are being raised about valid uses of I.Q. tests. For example, many schools use I.Q. tests to determine mental retardation. Recently, a California hospital compared over a five-year period the incidence of children who were mentally retarded in the community with those so labeled in schools in the community. They found that based on norm-referenced I.Q. tests, 100% more blacks and 300% more Chicanos were classified as mentally retarded in the schools than were in the community population (Mercer, 1976).

Three logical conclusions come to mind as to why there were such gross mistakes made in classifying these students: (1) the tests were misused by teachers and counselors; (2) the tests were not normed to include a true representation of black and/or Chicano norms; or (3) the tests both were misused and had been invalidly normed.

In my judgment, one of the reasons norm-referenced tests are invalidly normed is the use of item analysis in validation of referenced tests. (See #5 on Chart A.) To use item analysis to validate norm-referenced tests, the test developer gathers some questions that experts consider to be representative of whatever the test covers. Then he gives these questions to (ideally) a random sample of the population represented by the test. Upon scoring the answers, the test-maker divides the scores into an upper 50% and a lower 50%. He then either throws out or rewrites the questions that more students in the lower 50% answered correctly than did the students in the upper 50%.

We are beginning to realize that certain advantaged groups consistently score higher on typical norm-referenced tests than do other, less advantaged groups. The groups who tend to score higher are, at a minimum, the groups for whom the tests were truly normed.

If, in these tests, questions are purposefully included because they are easier for some groups to answer correctly, while other questions are thrown out purposefully which are easier for other groups to answer correctly, then I suggest that vocational educators need to understand this fact and how it can affect vocational educators and their students.

If the use of norm-referenced tests gives an abnormal picture of some students, how can vocational educators accurately individualize the instruction of these students? This problem is so serious that at least one state has outlawed the use of group I.Q. tests (Newsweek, March 22, 1976). One of the reasons for this problem is that group I.Q. tests are normed for people who are competent in reading. If a person who takes these tests is low in reading skill, then he stands a good chance of being mislabeled as having a low I.Q., instead of being classified merely as needing to gain reading competency. How many students are being misclassified (and, therefore, mistaught) in all of education? Do not the constituents of vocational education many times lack reading competencies? We need to know how many are misclassified because of the use of all norm-referenced tests--we need to know not only the uses of norm-referenced tests, but also their misuses.

Criterion-referenced tests do not have the problems with norm validation that are prevalent in norm-referenced tests (Chart A). Valid criteria are the problem with criterion-referenced tests. What criteria are valid to use to ascertain what accomplishment? For example, in vocational education, these criteria may range from easily judged criteria such as typing speed to difficult to judge criteria such as ability to sell. If one of vocational education's missions is to aid students in learning salable skills, then salable skills are some criteria or performance standards to use in vocational education criterion-referenced tests.

A problem does arise with criterion-referenced tests concerning the assumption of the validity of the criteria. Are the criteria used in such tests valid ones? For instance, if criteria are based on their "salability," how does one determine if they are salable? This kind of validity is difficult. One of the ways to determine if performance standards are salable is to compare them with skills that have been sold. For a better explanation of criterion validity in vocational education, examine the method that the Vocational-Technical Education Consortium of the States uses. Suffice it to say that criterion validity is difficult to achieve.

In summary, if the mission of vocational education is to help provide students with salable skills, then of the three available methods of evaluating vocational education outputs, criterion-referenced tests seem to be most adequate.

Evaluation of Outputs and/or Outcomes

The issue here is whether or not vocational education can realistically be evaluated based on both outputs and outcomes, either outputs or outcomes, or neither outputs nor outcomes. For comparative purposes, I shall first discuss outputs, then outcomes, and then compare the uses of the two.

I refer to output as a product of vocational education--a person who comes out of vocational education. Each output of vocational education takes with him or her competencies in dealing with data, people, and things. One can (and some do) argue that the skills learned by the person make up the output of vocational education; they argue that output refers to skills or competencies rather than to people. Admittedly, life would be simpler for all educators if skills could be easily separated from the people who possess them. However, individuals' attitudes and abilities may affect their competency development more so than can vocational education. Also, other courses in school affect student attitudes and abilities and, therefore, affect student competencies. Innumerable other factors also affect student competencies--for example, Chart B, meant to be by no means inclusive, shows some of the interrelationships of factors that affect competencies.

If the relationships in Chart B are even slightly representative of reality, we can see the difficulties of validity separating some competencies from others, and competencies from persons, and persons from their environment. If the various factors shown in Chart B do affect and are affected by each other, then how can we evaluate one factor without taking into account other factors? For example, what if students in vocational education have very high reading aptitudes, and other students do not? If this should be so, vocational education students have the possibility of developing more competencies that depend on reading than do the other students. Suppose we evaluate these outputs of vocational education in terms of reading-dependent competencies. Can we realistically state that the reason for which a person in vocational education has developed these reading competencies is solely his vocational education? Of course not. Thus,

CHART B

Interrelationships of Some Factors Which Affect Competencies

Student attitudes + student aptitudes	⟷	* student competencies**
Teacher attitudes + teacher aptitudes	⟷	teacher competencies
School attitudes + school aptitudes	⟷	school competencies
Occupational attitudes + occupational aptitudes	⟷	occupational competencies
Business attitudes + business aptitudes	⟷	business competencies
Social attitudes + social aptitudes	⟷	social competencies
Political attitudes + political aptitudes	⟷	political competencies
Economic attitudes + economic aptitudes	⟷	economic competencies

*⟷ means "affect and are affected by."

**Competency means "data skills + people skills + things skills."

the question should be: What part can vocational education play in the development of competencies, given that other factors also affect competency development? In essence, we need to ask, "For what influences on competency development can vocational education be held responsible, given other influences on students?"

Competencies are composed of skills--skills in regard to data, people, and things. When talking about skills, some people automatically assume that vocational skills mean salable skills, skills that can and will be sold in the job market. I question this assumption. First, there is a difference between "can be sold" and "will be sold." Skills that "can be sold" apply to the skills (some of which can be learned in vocational education) which contribute to occupational competencies. One can determine skills that can be sold through comparing these skills with those used in the job market. However, if one refers to skills that "will be sold," one is referring not only to effects of vocational education but also to effects of other factors, including the job market. The job market itself is a prime determinant of what will be sold. I assume that vocational educators can significantly affect vocational education--sometimes not to the degree desired, but they can affect vocational education. However, I contend that vocational educators cannot really affect the job market. We need not delude ourselves into believing that we can. There are some references in the literature which state that one purpose of vocational education is to reduce unemployment

through correlating occupational competencies with employment opportunities. This purpose is fallacious, for the following reasons.

In a competitive economy, when employment goes up in one area, it may go down in another. According to governmental policy, full employment is defined as 4% unemployment. (If the job market goes below 4% unemployment, problems supposedly are created by inflation.) Unemployment, thus, is used as an economic tool by government and business to keep inflation down. At present, also supposedly because of inflation, the official national unemployment rate is above 7%. It would thus seem that if the main way to increase employment is to reduce inflation, then the main way to reduce unemployment is to reduce inflation. Can vocational education do that?

If there were zero per cent unemployment, then the job market would be able to create demand for all the outputs of vocational and other education. Every person would have a skill that could be and, if desired, would be sold. Presently, with at least 7% unemployment (which is a conservative figure), the national job market is, at best, only 93% of its potential size. Using the size of the job market as a factor in determining skill salability, the average person has a maximum employment potential of 93%. In some areas this potential is even less; unemployment is much higher and employment potential based on job market size is much lower. It seems, then, that the true salability of skills depends not only on occupational competencies but also, at a minimum, on employment potential--the potential to sell skills in the job market.

Additionally, other factors not related to occupational competency significantly interfere with the ability of vocational education outputs to sell their skills in the job market. Some recent studies conclude that factors other than occupational competency and the size of the job market may be deciding factors in hiring and promotion. These factors range from race and sex to church attendance (Quinn et al., 1968). Which of these outside factors can vocational education influence? On which of them can vocational education be validly evaluated? Any? Some? None? We need to find out.

Within this discussion I have been functioning under two assumptions which now need to be made known: (1) no entity should be evaluated on something over which it has neither authority nor influence, and (2) an entity should be evaluated to the degree to which it has authority and influence. Can vocational education exert any significant authority and influence over how well its outputs can sell their skills in the job market? If so, how? We can affect the competency of our students; but if one accepts the earlier description of other factors affecting competency development as well, one realizes that many factors other than vocational training (such as reading ability) affect competencies as well as the degree to which occupational competencies compare with those already in the job market. If we accept the conclusions of recent studies, we realize that occupational competency may not be the deciding factor in hiring and promotion.

Thus, we come to the issue of outcomes. "Outcomes" here refers to the interaction of vocational education's products with the environment, especially with that part of the environment containing jobs. To what degree can vocational education influence the outcomes of its products, be they graduates, dropouts, or pushouts? To what extent can vocational education influence hiring (placement) and promotion? If we can have authority over and can influence hiring and promotion, then we should be evaluated on these outcomes. However, we also would need to evaluate the degree to which we can affect these outcomes. For example, if the unemployment rate for 16-19-year-olds in one area is 20%, there is room in this job market for only 80% of those 16-19-year-olds seeking jobs. If our outputs in this job market area are 16 to 19 years old, and if they are typical of that age group, can we provide them <u>all</u> with skills that can be sold? Given our true authority and influence, can we provide them with skills that transcend the other factors which authority and influence perhaps greater than that of vocational education?

Ideally, if the job market were large enough and if people were hired and promoted based only on their occupational competencies, theoretically, at least, vocational education could be evaluated somewhat more realistically regarding the outcomes of its products. Until, however, factors other than the limited authority and influence of vocational education are taken into consideration, the evaluation of vocational education's outputs and outcomes will continue to be unrealistic.

We have before us a goal: To determine what vocational education is being evaluated on and what vocational education should be evaluated on, and then to determine what vocational education can validly be evaluated on. When we go somewhere, we need to know not only where we are and where we should go, but also where we <u>can</u> go, given other limiting factors.

Evaluation of Efficiency and Effectiveness

The issue to be discussed here is adequacy of evaluating organizational efficiency and effectiveness; that is, whether evaluation of only the efficiency and effectiveness of an organization constitutes adequate evaluation.

Within near history, the two basic evaluation frameworks have been efficiency and effectiveness. When one looks at an organization through the framework of efficiency, he is comparing how much goes into the organization (input) versus how much comes out (output). For example, one might compare how much money goes into producing a good or a service versus how much money comes out of selling the good or service. If theoutput is greater than the input, profit results; if input is greater than output, a loss results. The ideal of efficiency is to have the least going in and the most coming out.

Effectiveness is complementary to efficiency. Effectiveness refers to how much the organization actually accomplishes versus how much the organization wants to accomplish. That is, effectiveness refers to the degree to which organizational goals are attained.

For business and industry, the evaluation of efficiency has been basically simple. All other things being equal, the efficiency of business and industry can be judged by comparing the amount received with the amount spent. In other words, business and industry can use monetary profits or losses as a criterion of efficiency, and increased efficiency should yield increased profits. Effectiveness, however, has played second fiddle to efficiency. It has usually been assumed that the only goal of business is to make profits, so the question has largely been: To what degree has the business organization attained its profit goal? But, as you can see, if one looks at any organization in terms of efficiency and effectiveness only in terms of profits, one is placing blinders on himself about other possible social, political, and economic effects of the organization.

Adequate evaluation of efficiency in vocational education is nearly impossible. This approach to evaluation, usually termed cost-benefit evaluation, has the following limitations.

Evaluators cannot gauge very accurately the money that goes through the vocational education system. There is also little evidence that other inputs to vocational education--students, teachers, time, etc.--can be measured accurately. Likewise, at the other end of the vocational education system, what outputs can be accurately gauged which can be reliably compared with inputs? Money? Number of students? Competencies? Further, what quantities or qualities of vocational education's outputs can be realistically judged to be a product of the vocational education system and not of other outside influences (if only socioeconomic level)?

Let's use "students-in" compared with "students-out" as a measure of efficiency. The most efficient vocational education organization is the one which takes in X number of students at one end of the system and exits the same number at the other end. There are many ways to fake or guarantee efficiency based on numbers of students. One is to take in only those students who have a high potential for graduating. "Appropriate" entrance requirements can be set so as to admit only guaranteed graduates. Another is to guarantee a diploma to every student, regardless of what the student accomplishes.

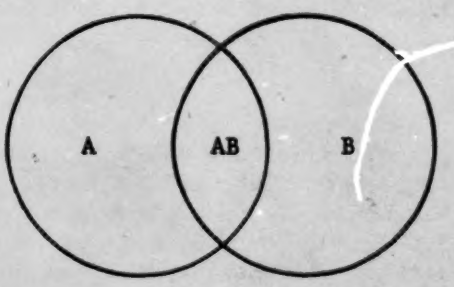
But we must consider reality. Vocational education cannot even attempt to take in only guaranteed students. Only such schools as M.I.T. may afford this luxury. Taking into consideration the constituents of vocational education who need help the most, such a practice would, in fact, be not only unlawful in some cases, but probably unethical in the majority of cases.

Thus, criteria upon which to evaluate efficiency include not only the number and types of students, but also, at least, the context of these inputs and the means by which individuals are processed through the organization. Context, inputs, and processes also affect and are affected by what an organization wishes to accomplish. For example, one of M.I.T.'s goals may realistically be to produce engineers who place in the top 5% of all who take a national professional engineering exam. Their efficiency and the degree to which they accomplish their goal depend on context, input,

and processes. If they do not have sufficient equipment, qualified students, and sound teaching procedures, they probably will have difficulty reaching their respective goals.

We have come to the evaluation of effectiveness in vocational education; that is, the degree to which desired goals are accomplished. The closer an organization comes to reaching its goals, the more it will be evaluated as being effective. But what if the only goal of the organization is merely to exist? If the organization does exist, then, based on its goal, it is an effective organization. This organization performed well enough to exist, but is existence enough?

In order to determine if goals are sufficient, goals should be examined with regard to the existing context of the organization, inputs into the organization, and processes used by the organization. The minimum goal of any kind of organization is, of course, to exist; but, when considering the context, inputs, and processes of vocational education, its goal cannot reasonably be to place students in the top 5% on a national engineering exam. Vocational education's goals are obviously somewhere between the two extremes. If the goals are not realistic, then the evaluation of effectiveness will not, or rather cannot, be realistic. Following is a pictorial representation of the convergence of realistic goals with contemporary evaluation:



A = realistic goals
B = contemporary evaluation
AB = contemporary evaluation with realistic goals
B-AB = contemporary evaluation without realistic goals

Policy Evaluation

Policy is defined in Webster's Seventh New Collegiate Dictionary as "a definite course or method of action selected from among alternatives and in light of given conditions to guide and determine present and future decisions." Policy thus encompasses more than the persons who occupy positions in an organization. People and positions sometimes change rapidly. Policy is a more constant and consistent link between an organization's past and present and between its present and future than are people and positions.

There are both written and unwritten policies. Unwritten policies may be analogous to norms of an organization--norms which influence present and future decisions. Written policy expresses authoritative norms in writing, as well as setting parameters for what other norms should be. In order to be authoritative, the written policy of an organization should fall within the zones of acceptance of the organization's constituencies. Policy as a communication of authority. However, Chester Barnard hypothesizes that (Barnard, 1968):

> a [policy] will be authoritative only if it possesses the following characteristics:
>
> (a) the person to whom it is addressed can and does understand it;
> (b) he believes it to be compatible with the purposes of the organization and his personal interests as a whole;
> (c) he is physically and mentally able to comply.

Good policy is diffused throughout the organization and applies to all within the organization. Therefore, policy statements are not specific instructions for specific individuals. According to David Bell (1975),

> Explicit instructions . . . have two liabilities:
>
> (a) they outline by implication what poeple do not have to do;
> (b) they apply only to those to whom they are specifically sent, thereby relieving nonrecipients of the responsibility of compliance.

Written policy is an aid to the evaluator; he can know more generally what he should evaluate and can put the evaluation into a more realistic context than he could if he had only unwritten norms to use. However, in many cases, the more that norms are understood and accepted (or the more that authority is absolute), the less there is of written policy. Unwritten policy has utility to persons in positions of authority. If there is only unwritten policy about an area, the persons in authority will be the only ones who "know" ultimate answers. If one cannot refer to written policy, one must refer to persons in positions of authority. Thus, without written policy, arbitrary and capricious judgments may "legitimately" be made by persons in authority positions.

Another way to look at policy is to determine the level of one policy in regard to other levels of a hierarchy, ultimately to the highest level. According to Dale Mann (1975):

> In a sense it is possible to see policy issues as a middle stratum; above policy problems are macro-societal problems and below policy problems are operational ones. Thus, [policies] are not concerned with the day-to-day tasks of managing a classroom, operating school buildings, or administering a district. To the extent that these latter

activities are concerned with tactical details, administrative routines, the nuts and bolts of housekeeping management, they will be below the threshhold defined as "policy" here.

Thus, policies can be viewed as a means of transaction between macro-societal problems and operational ones. The policies of vocational education would be formal, authoritative links going from macro-societal problems through vocational education to operational problems.

So, what policies are there for vocational educators to follow? How are vocational educators to know what they are responsible for and, therefore, what they are not responsible for? The answer to this question is the issue at hand. Rationally, when one evaluates policy at one level, he should also be able to compare how well this policy fits into the policies at the highest level. Ideally, all levels of policy are congruent, and one can evaluate a policy based on its "goodness of fit" with policies of other levels.

Other means of policy evaluation do exist. According to the National Advisory Council on the Education of Disadvantaged Children (1976), the following are

requirements of effective evaluation that are policy level rather than merely performance level:

1. The evaluation must test the program's basic assumptions. No evaluation of a policy level order can be judged effective unless constructed around a test of the assumptions behind the interventions a law mandates. It is the task of evaluation to test assumptions about what works and why. It should help identify correct assumptions [underlining added]. So, evaluation design must produce an organized analysis to produce the data from which can be extrapolated quantitative and qualitative insights, validating, refuting, or modifying the operative set of assumptions behind law and practice.

2. Evaluation must proceed from clear definition of the problem addressed and the program intervention made. Definition is vital. Lacking it, one cannot be certain that what is being measured is, in fact, the response to the original problem [underlining added]. It might merely be the program rationale which, like a protective coat, keeps many programs "warm" even after well-distancing original purposes. One also cannot be sure, lacking clear definition, that the measurements which evaluation arrives at really relate to or provide relevant feedback in terms of Congressional intent, or represent significant insights.

3. Everyone who is key must want the evaluation and work at making it work, whether Congressional Committee(s), OMB, OE/HEW or /NIE, or local level education administrators.

4. Regrettably, there can be no objective, meaningful evaluation of education cost effectiveness, because there is no accurately testable correlation between expenditures to produce the intervention and program implementation, program outcome or performance and, finally, program impact.

Unfortunately, vocational educators have few levels of policy to use as criteria for judging "goodness of fit" of their own policies. There is a dearth of both local and State Board policies. Perhaps even more importantly, there is no national policy for vocational education which could be used by evaluators as a reference point or as a criterion for evaluative purposes.

According to the National Advisory Council on Vocational Education (1976),

> All vocational education's changes have taken place as a random response to circumstances--not as the orderly implementation of a comprehensive national policy for vocational education.
>
> The consequences of the lack of a national policy for vocational education are becoming more and more severe. It has meant that the scope of responsibilities of the vocational education establishment has never been clearly defined. . . .
>
> It has meant that vocational education is exceptionally vulnerable to attack from its critics--not because it is ineffective, but because there is no national policy against which its effectiveness can be measured or, in other words, evaluated [underlining added].

Thus, while evaluation in vocational education has focused on performance, such performance is not adequately covered by policy. Policy constitutes criteria of values. In the absence of values criteria, can evaluators of vocational education logically furnish information to decision-makers about vocational education upon which the decision-makers can realistically judge the worth of vocational education? In my judgment, we cannot presently evaluate vocational education accurately, not only because of the inadequacies in the state of the art in, and use of, evaluative techniques, but also, and more importantly, because we have no top level, no national policy context, within which to put any evaluation of vocational education.

I urge all to read the report of the National Advisory Council on Vocational Education entitled "A Call for a National Policy on Vocational Education." As you continue learning to evaluate, learn not only how to evaluate performance, but also to evaluate one of the controlling mechanisms of authority--the particular controlling mechanism called policy.

References

Angoff, William N. "Scales with Nonmeaningful Origins and Units of Measurement," Educational and Psychological Measurement, 22(1), 1962, 27-34.

"A Ban on I.Q. Tests?" Newsweek, March 22, 1976, 49.

Bell, David U. J. Power, Influence, and Authority. New York: Oxford University Press, 1975, 96.

Bernard, Chester. The Functions of the Executive. Cambridge, Massachusetts: The Harvard University Press, 1968, 163.

Cox, Richard C. "Evaluative Aspects of Criterion-Referenced Measures." In W. J. Popham (ed.), Criterion-Referenced Measurement: An Introduction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1971, 67-75. (ED 038 679)

Cox, Richard C., and Julie S. Vargas. "A Comparison of Item Selection Techniques for Norm-Referenced and Criterion-Referenced Tests." Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, Illinois, February, 1966. (ED 010 517)

Davis, Frederick B. "Criterion-Referenced Tests." In the proceedings of the Thirty-Fifth Annual Conference of the Educational Records Bureau, Greenwich, Connecticut, 1972, 40-42.

Flannagan, John C. "Units, Scores, and Norms." In E. F. Lindquist (ed.), Educational Measurement. Washington, D. C.: American Council on Education, 1951, 695-763.

Freeman, Richard, and J. Herbert Holloman. "The Declining Value of College Going," Change, September, 1975.

Gagne, R. M. "Some Notes on Criterion-Referenced Measurement." Mimeograph. Florida State University, December, 1969.

Glaser, R. "Instructional Technology and the Measurement of Learning Outcomes: Some Questions," American Psychologist, 18, 1963, 519-521.

Klein, S. P., and J. Kosecoff. Issues and Procedures in the Development of Criterion-Referenced Tests. Report No. 26. Princeton, New Jersey: ERIC Clearinghouse on Tests, Measurements and Evaluation.

Mann, Dale. Policy Decision-Making in Education. New York: Teachers College Press, 1975, 5.

Mercer, Jane R. "System of Multicultural Pluralistic Assessment (SOMPA)." Presentation at the Council for Exceptional Children Conference, New Orleans, Louisiana, February 12, 1976.

CE 008 559

National Advisory Council on the Education of Disadvantaged Children. "Can We Afford Deficient Evaluations?" Washington, D. C.: January 23, 1976, 7-8.

National Advisory Council on Vocational Education. "A Call for a National Policy on Vocational Education." Washington, D. C.: December 15, 1975, 3-4.

Popham, W. J., and T. R. Husek. "Implications of Criterion-Referenced Measurement," Journal of Educational Measurement, 6(1), 1969, 1-9.

Quinn, Robert P., et al. The Decision to Discriminate: A Study of Executive Selection. Ann Arbor, Michigan: Institute for Social Research, 1968, 16.

Shoemaker, D. M. "Criterion-Referenced Measurement Revisited," Educational Technology, March, 1971.

Thorndike, E. L. "Original Tendencies as Ends: Emulation in the Case of School 'Marks.'" In E. L. Thorndike, Educational Psychology. New York: Teachers College Press, 1913, 286-289.

Thorndike, Robert L., and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. New York: John Wiley and Sons, Inc., 1969, 326-327.

Ward, J. "On the Concept of Criterion-Referenced Measurement," British Journal of Educational Psychology, 40, 1970, 314-323.